Biochem. J. (1986) 240, 357–360 (Printed in Great Britain)

A suitable parameterization of the Michaelis–Menten enzyme reaction

D. A. RATKOWSKY
CSIRO Division of Mathematics and Statistics, Tasmanian Regional Laboratory, Stowell Avenue, Hobart, Tasmania 7000, Australia

It is shown here that a suitable form for estimation and inference using the Michaelis–Menten [(1913) Biochem Z. **49**, 333–369] model for simple enzymic reactions is one in which the two parameters appear in the denominator of the equation. In this form, convergence to the least-squares estimates using the Gauss–Newton method [see Kennedy & Gentle (1980) Statistical Computing, Marcel Dekker, New York] is virtually ensured, or, as the model in this form is a member of the class of 'generalized linear models', it may be fitted by packages such as those of Rothamsted Experimental Station [(1977) GENSTAT (A General Statistical Program), Rothamsted Experimental Station, Harpenden] and the Numerical Algorithms Group [(1978) GLIM (Generalised Linear Interactive Modelling), Numerical Algorithms Group, Oxford]. Furthermore, the parameters-in-denominator principle is readily extended to more complicated catalytic models. With all parameters in the denominator, the least-squares estimators are close to being unbiased and normally distributed, whereas severe bias and non-normality may result from use of the standard formulations.

INTRODUCTION

One of the most commonly used formulations for modelling enzyme kinetic reactions, where a single substrate forms a complex with the enzyme, is the Michaelis–Menten (1913) model expressed as:

$$v = \frac{V_{max.}[S]}{[S]+K_m} \quad (1)$$

where v is the velocity of the reaction, [S] is the substrate concentration, and $V_{max.}$ and K_m are parameters to be estimated. This is the equation of a rectangular hyperbola, with $V_{max.}$ representing the maximum velocity theoretically obtainable and K_m being the value of [S] at which the velocity is half the maximum velocity.

If the errors in v about the regression model can be assumed to be normally distributed and of the same magnitude for all substrate concentrations, [S], then the appropriate procedure is to use non-linear least squares. An exact solution can be obtained using the Gauss–Newton method [see Chapter 10 of Kennedy & Gentle (1980)], which requires good initial estimates of the parameters. Nevertheless, it is important to realize that, unlike linear-regression models, where the least-squares estimators are unbiased, normally distributed, minimum-variance estimators, the estimators of the parameters in non-linear models may be badly biased, non-normally distributed and have variances greatly in excess of the minimum possible variance. This bias exists because the regression model is non-linear in its parameters. However, as the sample size increases toward infinity, the bias diminishes, the distribution of the estimator becomes more normal and the excess variance decreases, thereby approaching more and more closely the condition for a linear model. Some non-linear regression models approach the large-sample behaviour even in small samples; I termed such models 'close-to-linear' (Ratkowsky, 1983) and advocated searching for, and identifying, such models for practical use.

Eqn. (1) may be reparameterized by putting the parameters in the denominator, that is, by rewriting it as follows:

$$v = \frac{[S]}{\theta_1[S]+\theta_2} \quad (2)$$

Here, the parameter θ_1 is simply $1/V_{max.}$, the reciprocal of the maximum velocity, and parameter θ_2 is $K_m/V_{max.}$, the ratio of the two parameters in eqn. (1). In a study of general formulations for modelling catalytic chemical reactions, I (Ratkowsky, 1985) concluded that putting the parameters in the denominator was the way of obtaining a close-to-linear model for that class of problems. Eqn. (2) is an example of that class, the only difference being that the reaction is biochemical instead of chemical.

STATISTICAL METHODS

I (Ratkowsky, 1983) described methodology for examining the statistical properties of the least-squares estimators of the parameters in non-linear regression models. Among the various available measures are the curvature measures of intrinsic (IN) and parameter-effects (PE) non-linearities of Bates & Watts (1980), the bias measure of Box (1971) and the asymmetry measure of non-linearity of Lowry & Morton (1983). IN measures the curvature of the solution locus and should be close to zero if the solution locus is acceptably straight. Reparameterization does not alter the solution locus, so all parameterizations of the same basic model have the same IN values. PE measures the straightness, parallelism and equi-spacedness of the parameter lines on the solution locus (actually on the tangent plane to the solution locus) and should be close to zero for a close-to-linear model. The bias measure of Box (1971) quantifies the extent of the bias in the estimates of the parameters. The Lowry & Morton (1983) λ-values give a measure of the non-linear behaviour of the estimator

for each parameter separately. A rule of thumb based on extensive use of this measure is that if $\lambda < 0.01$, the estimator is close-to-linear, but larger values indicate increasing skewness in the distribution of the estimator. If $\lambda > 0.05$, the skewness is very perceptible if histograms of the estimates are drawn after carrying out the simulation study I described (Ratkowsky, 1983).

To obtain initial estimates of θ_1 and θ_2, eqn. (2) can be rewritten in the following form:

$$\frac{[S]}{v} = \theta_1[S] + \theta_2$$

so that by regressing $[S]/v$ versus $[S]$ or graphically plotting $[S]/v$ versus $[S]$ [the Hanes (1932) plot], initial estimates are obtained that can then be used in conjunction with the Gauss–Newton method to determine the least-squares estimates, $\hat{\theta}_1$ and $\hat{\theta}_2$, of θ_1 and θ_2 respectively. Watts (1981) graphically demonstrated why the linear approximation method first proposed by Gauss and used in the Gauss–Newton method ensures rapid convergence for a close-to-linear model. This is because the solution locus is closely approximated by the tangent plane and because the uniform co-ordinate system of straight, parallel and equi-spaced lines on the tangent plane closely approximates the positioning and spacing of parameter curves on the solution locus. If the model were exactly linear, convergence would be achieved in a single step from any starting point [see Ratkowsky (1983) for a proof of this]. A close-to-linear model requires more than one step, but convergence will be rapid.

Once $\hat{\theta}_1$ and $\hat{\theta}_2$ have been obtained, the least-squares estimates, $\hat{V}_{max.}$ and $\hat{K}_m$, of the parameters $V_{max.}$ and K_m of eqn. (1) are readily obtained by direct substitution into the formulae $\hat{V}_{max.} = 1/\hat{\theta}_1$ and $\hat{K}_m = \hat{\theta}_2/\hat{\theta}_1$. An estimate of variance (σ^2) of $\hat{K}_m$ may be obtained from:

$$\hat{\sigma}_{K_m}^2 = \hat{K}_m^2[(\hat{\sigma}_1^2/\hat{\theta}_1^2) + (\hat{\sigma}_2^2/\hat{\theta}_2^2) - 2(\sigma_{12}/\theta_1\theta_2)]$$

where $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ and $\hat{\sigma}_{12}$ are estimates of the variance of $\hat{\theta}_1$, the variance of θ_2 and the covariance of θ_1 and θ_2 respectively. Approx. 95% confidence limits for K_m are then obtained from $\hat{K}_m \pm 2\hat{\sigma}_{K_m}$.'

A total of eleven data sets are considered here. The data sets are taken from the following sources: (1) Bates & Watts, 1980, p.10; (2) Wilkinson, 1961, p. 329; (3) Wong, 1975, p. 245; (4)–(6) Bliss, 1970, p. 101, 103 and 113; (7)–(11) Apitz *et al.*, 1971, p. 361.

RESULTS

Table 1 presents results for IN and PE for eqns. (1) and (2). Data sets having similar values of these measures have been grouped, and average values are presented. As IN is unaltered by a reparameterization, its values are the same for both models. For all data sets, IN is adequately low, and only for data set 10 does IN approach significance. PE is higher in every case of eqn. (1) than for eqn. (2), being significantly high for six of the data sets compared with only one significant value for eqn. (2).

Table 2 presents values of the bias in each of the parameter estimates, calculated by using the formula of Box (1971). In every case, the magnitude of the bias in the parameter estimates of eqn. (1) is greater than the bias in the estimates of eqn. (2), being in excess of 20% for both parameters for data set 10 for eqn. (1) compared with less than 2% for eqn. (2).

Table 1. Intrinsic (IN) and parameter-effects (PE) non-linearity measures of Bates & Watts (1980)

Average values are given for data sets having similar values of IN and PE. An asterisk (*) indicates significance at the 5% level.

	Non-linearity measure		
	IN	PE	
Data set	Eqns. (1) and (2)	Eqn. (1)	Eqn. (2)
4,5,6,8	0.021	0.078	0.039
2,7	0.046	0.182	0.078
11	0.044	0.301*	0.100
1,3,9	0.077	0.560*	0.171
10	0.147	1.619*	0.291*

Table 2. Bias in each parameter estimate, expressed as a percentage of the estimate, calculated by using the formula of Box (1971)

Average values are given for data sets having similar values of bias.

Data set	Parameter	Bias in estimates of Eqn. (1) (%)	Parameter	Bias in estimates of Eqn. (2) (%)
4,5,6,8	$V_{max.}$	0.042	θ_1	0.002
	K_m	0.126	θ_2	0.033
2,7	$V_{max.}$	0.248	θ_1	−0.024
	K_m	0.635	θ_2	0.133
11	$V_{max.}$	0.809	θ_1	−0.070
	K_m	1.496	θ_2	0.184
1,3,9	$V_{max.}$	2.908	θ_1	−0.241
	K_m	5.206	θ_2	0.581
10	$V_{max.}$	28.8	θ_1	−1.547
	K_m	42.3	θ_2	1.707

Table 3. Values of the asymmetry measures λ of Lowry and Morton (1983)

Average values are given for data sets having similar asymmetries. The values are also averaged over both parameters.

	λ-Value	
Data set	Eqn. (1)	Eqn. (2)
4,5,6,8	0.001	0.001
2,7	0.007	0.002
11	0.019	0.003
1,3,9	0.066	0.009
10	0.387	0.027

Table 3 presents values of the Lowry–Morton (1983) asymmetry measure, λ. For eqn. (1), λ exceeds 0.05 for four data sets, indicating perceptible skewness (and thus non-normality) in the distribution of the estimator. In no case does λ excess 0.05 for eqn. (2).

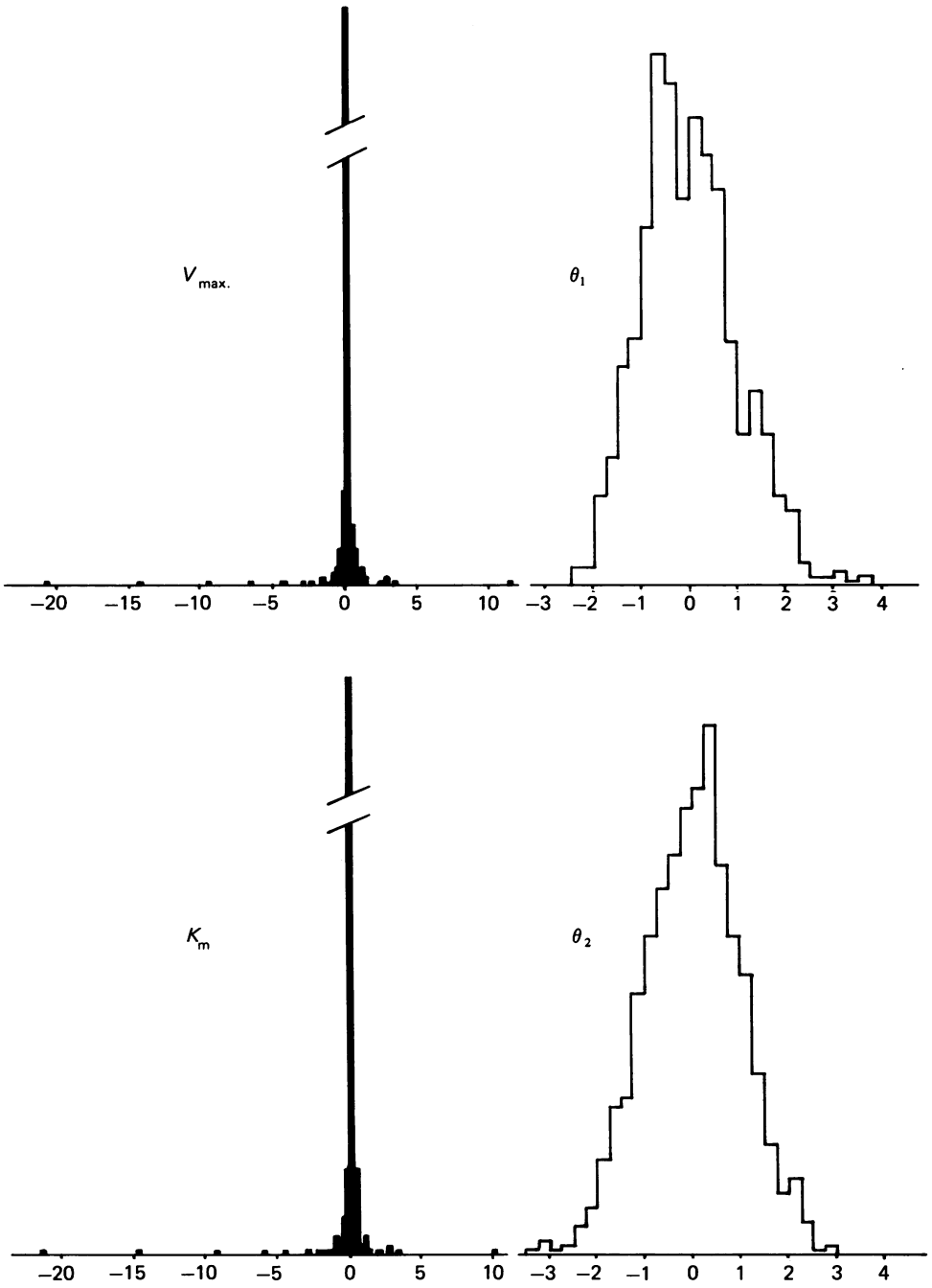

Fig. 1. Results of simulation study of 1000 trials for data set 10

Histograms of estimates of $V_{max.}$ and K_m are from eqn. (1), and histograms of estimates of θ_1 and θ_2 are from eqn. (2). In each case, the abscissa shows the distribution of the estimates standardized to have zero mean and unit variance. The ordinates represent the class frequencies.

Fig. 1 presents results of a simulation study of 1000 trials carried out for data set 10. The estimates of parameters θ_1 and θ_2 of eqn. (2) are both close to being normally distributed, which is what is expected from a close-to-linear model. In contrast, the estimates of parameters $V_{max.}$ and K_m exhibit extraordinary non-normal behaviour, with one estimate of each parameter appearing more than 20 S.D. away from the mean, and five other estimates being situated more than 5 S.D. from the mean.

DISCUSSION

The results presented in Tables 1, 2 and 3 show decisively that eqn. (1) is inferior to eqn. (2) in its statistical properties. For every data set, the parameters θ_1 and θ_2 of eqn. (2) have estimators which are closer to being unbiased, normally distributed, minimum variance estimators than the estimators of $V_{max.}$ and K_m in eqn. (1). The results of the simulation study of data set 10 presented in Fig. 1 indicate that estimates of K_m and

$V_{max.}$ can be obtained which are very different from their true values. This implies that the continued use of eqn. (1) may, for some data sets at least, lead to grossly misleading results. The estimators of the parameters of eqn. (2), however, always exhibits good statistical behaviour for each of the data sets.

The principle used to obtain eqn. (2) extends readily to more complicated models. One simply has to ensure that all parameters appear in the denominator of the expression. I showed this principle to produce a model that was close-to-linear in behaviour for catalytic reactions of the type used in the chemical process industries (Ratkowsky, 1985). For example, consider the following rate equation, which involves three concentrations (C_1, C_2 and C_3) and four parameters ($V_{max.}$, K_a, K_b and K_c):

$$v = \frac{V_{max.}K_aK_bC_1C_2}{1+K_aC_1+K_aK_bC_1C_2+K_aK_cC_1C_3}$$

Simply dividing numerator and denominator by $V_{max.}K_aK_b$ to give:

$$v = \frac{C_1C_2}{\theta_1+\theta_2C_1+\theta_3C_1C_2+\theta_4C_1C_3}$$

produces a model form in which the new parameters (θ_1, θ_2, θ_3 and θ_4) give rise to least-squares estimators with better statistical properties than the least-squares estimators of $V_{max.}$, K_a, K_b and K_c. Equations of the above form appear often in biochemical kinetics, and these equations are very amenable to rearrangement as follows:

$$\frac{C_1C_2}{v} = \theta_1+\theta_2C_1+\theta_3C_1C_2+\theta_4C_1C_3$$

so that multiple linear regression of C_1C_2/v on C_1, C_1C_2 and C_1C_3 produces estimates of θ_1, θ_2, θ_3 and θ_4 which serve as good initial estimates for the non-linear least-squares regression. Rapid convergence from the initial estimates to the least-squares estimates should occur when the Gauss–Newton method is used.

A further advantage of the 'parameters-in-denominator' models is that they conform to the specifications of generalized linear models (Nelder & Wedderburn, 1972) and therefore can be fitted by computer packages such as GENSTAT (Rothamsted Experimental Station, 1977) and GLIM (Numerical Algorithms Group, 1978). Eqn. (2) may be rewritten as:

$$v = \frac{1}{\theta_1+\theta_2(1/[S])}$$

in which case the denominator is a linear expression (called a 'linear predictor'). The linear predictor is related to the fitted values of the dependent variable via an 'inverse' or 'reciprocal' link function. When this model is fitted by using the GLIM or GENSTAT program, the same estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ are obtained as are obtained by using the Gauss–Newton method. Estimates of their standard errors are also the same. Similarly, more complex models such as eqn. (5) are converted into generalized linear models by dividing the numerator and denominator by C_1C_2. The resulting denominator is a linear predictor and is related to v via a reciprocal (or inverse) link. Users need not supply initial parameter estimates since these are generated internally by the GLIM and GENSTAT programs.

There will be times when the assumption of independent and identically distributed normal error will not be tenable. If the errors are normally distributed but have variances which depend upon the substrate concentration, then weighted regression needs to be employed. The appropriate weighting is to make the weights inversely proportional to the variance corresponding to a given substrate concentration. Hence, by using eqn. (2), one would carry out the non-linear regression by minimizing:

$$Q = \sum_{i=1}^{n} w_i[v_i-[S]_i/(\theta_1[S]_i+\theta_2)]^2$$

where n is the number of data points, $[S]_i$ and v_i are the substrate concentration and reaction velocity respectively for the ith data point, and w_i is the weight which will be chosen to be inversely proportional to the variance σ_i^2 corresponding to $[S]_i$. For the independent and identically distributed normal-error case discussed above, all w_i values are identical, so this term may be omitted from the expression.

Received 14 March 1986/11 August 1986; accepted 14 August 1986

REFERENCES

Apitz, R., Michelson, K., Shriver, K. & Cordes, E. H. (1971) Arch. Biochem. Biophys. **143**, 359–364

Bates, D. M. & Watts, D. G. (1980) J. R. Statist. Soc. **B42**, 1–25

Bliss, C. I. (1970) Statistics in Biology, McGraw-Hill, New York

Box, M. J. (1971) J. R. Statist. Soc. **B33**, 171–201

Hanes, C. S. (1932) Biochem. J. **26**, 1406–1421

Kennedy, W. J. & Gentle, J. E. (1980) Statistical Computing, Marcel Dekker, New York

Lowry, R. K. & Morton, R. (1983) Proc. 44th Sess. Int. Stat. Inst. Madrid, Contrib. Pap. **1**, 351–354

Michaelis, L. & Menten, M. L. (1913) Biochem. Z. **49**, 333–369

Nelder, J. A. & Wedderburn, R. W. M. (1972) J. R. Statist. Soc. **A135**, 370–384

Numerical Algorithms Group (1978) GLIM (Generalised Linear Interactive Modelling), Numerical Algorithms Group, Oxford

Ratkowsky, D. A. (1983) Nonlinear Regression Modeling: a Unified Practical Approach, Marcel Dekker, New York

Ratkowsky, D. A. (1985) Chem. Eng. Sci. **40**, 1623–1628

Rothamsted Experimental Station (1977) GENSTAT (A General Statistical Program), Rothamsted Experimental Station, Harpenden

Watts, D. G. (1981) in Kinetic Data Analysis: Design and Analysis of Enzyme and Pharmacokinetic Experiments (Endrenyi, L., ed.), pp. 1–24, Plenum, New York

Wilkinson, G. N. (1961) Biochem. J. **80**, 324–332

Wong, J. T.-F. (1975) Kinetics of Enzyme Mechanisms, Academic Press, London